UDC 551.509.5:551.509.313:551.509.333

EVALUATION OF 500-MILLIBAR DAILY AND 5-DAY MEAN NUMERICAL PREDICTIONS

JAMES F. ANDREWS

National Meteorological Center, Weather Bureau, ESSA, Suitland, Md.

ABSTRACT

Selected monthly 500- and 700-mb mean-error patterns from the National Meteorological Center's extended numerical predictions to 144 hr are shown and discussed. The NMC model (barotropic after 48 hr) consistently underforecasts the amplitude of most troughs and ridges. The pattern of error is usually established by 48 hr with the magnitude increasing with time. Also presented are charts (for 1968) of seasonal mean 700-mb error for the 5-day averaged numerical forecasts centered 4 days in advance ($D+4$). The large-scale mean errors vary with season and are also highly correlated with the observed height anomaly.

Average error patterns for special extended runs of the NMC six-level primitive equation (PE) model are compared with those produced by the standard operational model. Except for a negative height bias in the PE model at low latitudes, especially in the warm season, there is very little difference at 500 mb.

The hemispheric 5-day mean upper level prognostic charts prepared from the NMC's numerical output have shown gradual improvement during recent years, as measured by RE skill, a statistic based on "reduction of error." Comparing $D+4$ mean forecasts made by direct averaging of the NMC numerical output with those generated by the Extended Forecast Division's (EFD) flow model shows that the former have maintained a consistent and gradually increasing advantage over the latter, except for the southeastern United States.

Removal of the more systematic large-scale numerical prediction errors, either by use of a running average error or, better yet, by circulation typing, can lead to improved circulation forecasts.

1. INTRODUCTION

The tendency for the NMC operational numerical prediction model to show rather systematic errors in many of the 500-mb forecasts has been recognized for some time. Numerous studies, for example, Martin (1958) and Dunn (1964) with the barotropic model, and more recently Fawcett (1969) with the baroclinic primitive equation (PE) model, have been made on this subject, but these have been concerned primarily with prognoses to 36–48 hr. As changes were made in the operational model, first by introduction of the three-level baroclinic model by Cressman (1963) in 1962, and the six-level baroclinic PE model by Shuman and Hovermale (1968) in June 1966, the forecasts were gradually extended in time.

These predictions have been automatically incorporatde into the EFD 5-day forecast routine where the basic procedure remains essentially the same as it was prior to the advent of numerical forecasting. Briefly, this procedure involves the preparation of a time-averaged midtropospheric circulation pattern for the period 2 to 6 days in the future. This forecast, known as the $D+4$ chart since it is centered 4 days in advance, is then interpreted in terms of anomalies of temperature and precipitation and is used in determining the general paths and genesis and decay of cyclones and anticyclones on the series of four daily prognostic sea-level charts from 72 to 144 hr.

Three basic 5-day mean charts incorporating the daily numerical predictions were used in preparation of the 5-day forecasts from 1958 to 1965, and have been described fully by Namias (1958). The first of these, known as the D_0, is centered on a forecast day and represents the initial state of the 5-day mean circulation. It is composed of the latest three observed 0000 GMT upper level maps and the 24- and 48-hr numerical forecasts. The $D+2$ (formerly called summation) is centered 2 days later and consists of the four daily numerical predictions to 96 hr and the latest observed 0000 GMT upper level chart. The third of these charts, the flow or basic current, is produced from a model developed in the EFD in order to supply a fundamental need for extension of the numerical forecasts. This model, based on empirical-physical reasoning, predicts a 5-day mean upper level chart centered 4 days in advance ($D+4$ period) using the $D+2$ chart as input.

Beginning in May 1965, daily 500-mb numerical prognoses were prepared out to 144 hr by barotropic extension of the three-level model (six-level PE model starting June 1966). These forecasts, run from 0000 GMT data on regular 5-day forecast days, Sunday, Tuesday, and Thursday, now made it possible to prepare a $D+4$ circulation forecast from daily prognoses by simply averaging the five predictions from day 2 to day 6.

As a result of recent advances in numerical prediction, the forecaster is able to devote less time to forecasting the average circulation and more time to the weather to be expected. Experienced forecasters can, by their evaluation of the daily numerical iterations, consideration of the mean approach, and use of various statistical techniques, produce a 5-day forecast superior to one prepared objectively.

One way of evaluating the daily 500-mb forecasts has been to prepare hemispheric charts of average monthly forecast height error for each of the 6 days. Each chart is the average algebraic error (bias) of the 12 to 14 forecasts at each time interval. These patterns are assumed to be representative of the entire month since the days chosen are very nearly evenly distributed.

The purpose of this report is to show and discuss representative patterns of daily (to 144 hr) and 5-day mean ($D+4$) numerical prediction error from the regular operational and six-level PE models, to relate these to the circulation, and to indicate possible changes in systematic error due to improvements in the model. A 2-yr comparison is also made, using seasonal reduction of error (RE) skill scores, of the relative performance of the D_0, $D+2$, $D+4$, and flow charts.

To properly evaluate extended period forecasts of the NMC operational model since 1965, one must keep in mind several dates when the most significant changes are thought to have occurred:

May 1965—Extension of the combined Cressman three-level barotropic model prognoses from 96 to 144 hr.

June 1966—Six-level primitive equation (PE) model placed in operation and run to 36 hr, with the barotropic model to 144 hr.

February 1967—Latent heat from precipitation incorporated into the PE model.

May–June 1967—Extension of the PE model from 36 to 48 hr.

July 1967—Solar radiative heating incorporated into the PE model.

September 1968—Rough mountains incorporated into the PE model over western North America.

2. MEAN NUMERICAL PREDICTION ERROR

MONTHLY ERROR PATTERNS FOR 48-, 96-, AND 144-HOUR FORECASTS

Figures 1 through 6 show monthly mean patterns of 500-mb error (tens of feet) for 48-, 96-, and 144-hr forecasts for July and January from July 1965 to January 1969. Also shown superimposed on the 96-hr error patterns are the observed mean contours for the corresponding series of forecasts. Observed mean contours for the 48- and 144-hr forecasts, would, of course, be very similar to those for the 96-hr predictions. All error patterns relate to the barotropic portion of the model, except for the 48-hr errors beginning July 1967 which are from the PE model. Error charts for the other months, commencing August 1965, are available from the Extended Forecast Division of the National Meteorological Center.

An examination of these patterns brings to light certain characteristics which are equally applicable to all seasons. The magnitude of the errors increase most rapidly to 96 hr, then increase slowly or remain about the same, with the pattern usually being established at 48 hr. The amplitude of most troughs and ridges is underforecast, as shown on the 96-hr patterns (figs. 2 and 5) as positive errors in troughs and negative errors in ridges. There is also a marked tendency for the centers to cluster about preferred geographical areas in certain seasons. This characteristic is also applicable to the monthly mean troughs and ridges which were observed in the same general areas shown by Stark (1965). Principal centers of positive error are found in or near the trough axis and just north of the jet maximum (as determined subjectively); major centers of negative error are frequently located in or near the ridge axis and just south of the jet maximum. It is also seen that at middle latitudes wave numbers 3 and 4 predominate during winter, while 4 and 5 predominate during summer. These general characteristics of the combined PE-barotropic model were also found to exist in 36-hr forecasts made from both the three-level and PE models (Fawcett 1969).

The large-scale error and circulation patterns for summer and winter, as represented by the charts for July and January, will now be discussed in more detail.

Summer—One of the most systematic errors in summer is the extensive area of negative error which appears in North America (figs. 1, 2, and 3). This error, which is largest in July, is well established at 48 hr (fig. 1) and is associated with the ridge usually found over the Rocky Mountains and Great Plains in summer. Note the strong tendency for two centers to be observed, one in the Northern Plains, the other in Alaska or northwest Canada. Much of this error appears to be related to too rapid eastward advection of vorticity from the trough in the eastern Pacific. Lack of heating in the model may also have contributed to the large error in the Northern Plains in July 1966, since this was the only July in which above-normal temperatures were observed over most of the Nation.

Daily mean error patterns for July 1969 (not shown) were not significantly different from those of earlier Julys (figs. 1, 2, and 3). Two centers of negative error were found over North America at 96 hr, one center of −340 ft in northwest Canada, another of −270 ft in the Northern Plains.

Much of the Atlantic and eastern Asia also tend toward a negative forecast bias. Centers of negative error in the eastern Atlantic troughs in July 1967 and July 1968 (fig. 2) are not readily explained.

A second area of persistent negative error in summer is the Mediterranean and North Africa. This error reaches its greatest magnitude in July and is in close correspondence to the center of maximum sensible heat transfer from surface to air (Budyko 1963). In July 1969, centers of negative error observed in North Africa had maximum values of −130 ft at 48 hr, −380 ft at 96 hr, and −430 ft at 144 hr. The slight decrease in this error at 48 hr (fig. 1) after July 1967 is probably due to the inclusion of solar radiative heating into the PE model (July 1967).

The only area of systematic positive error in summer is across higher latitudes of the Pacific. Here, the tendency for two centers to appear is also a reflection of the normal circulation, which has two troughs in the Pacific.

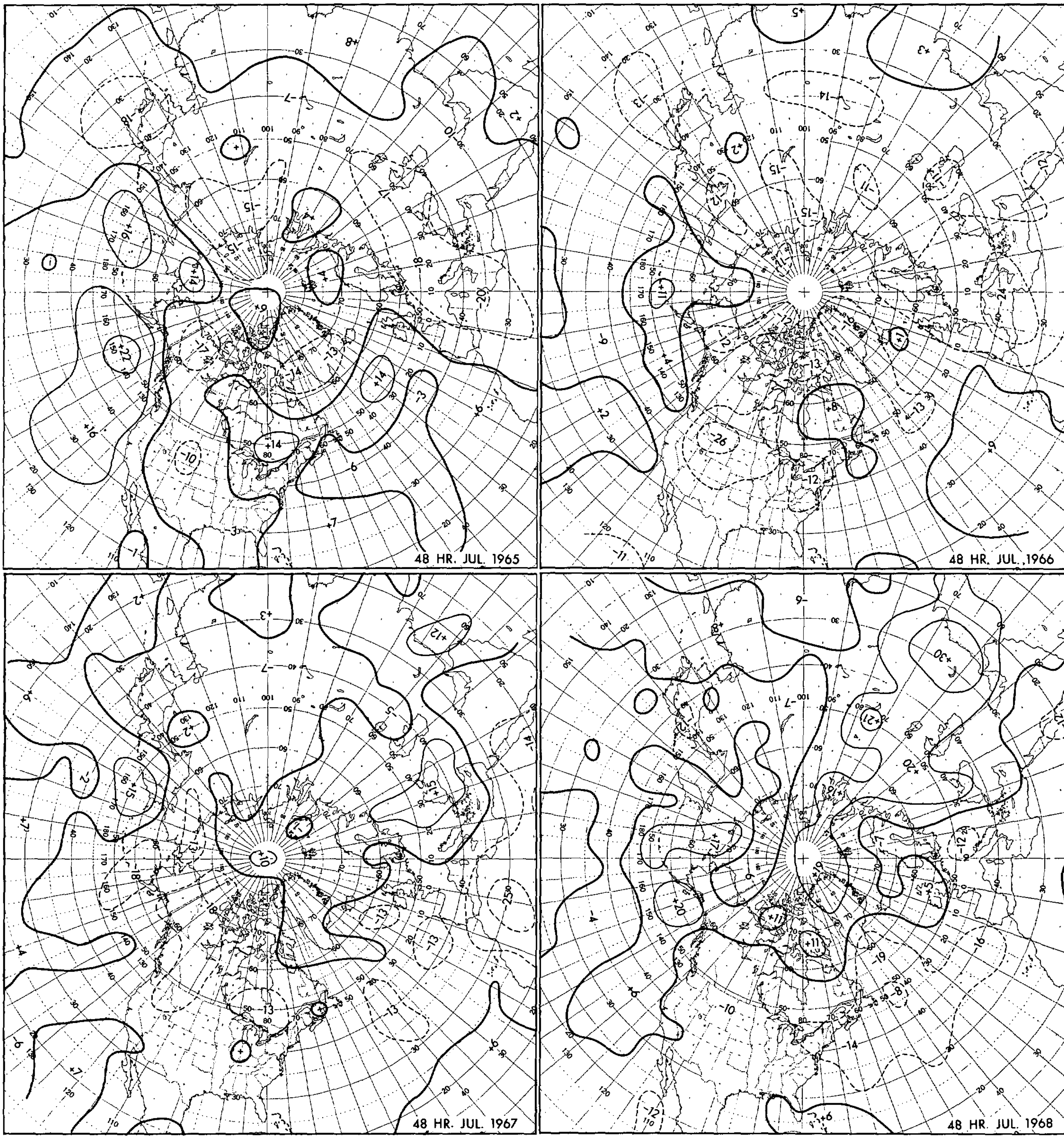

FIGURE 1.—Mean 500-mb 48-hr barotropic error for July 1965 and 1966 and 48-hr PE error for July 1967 and 1968. All error patterns in this and subsequent figures are in tens of feet and are valid at 0000 GMT.

Winter—The largest and most systematic error of the PE-barotropic model in winter is the positive error over the western Pacific (figs. 4, 5, and 6). It is related to a strong baroclinic zone combined with a major heat source which the model is incapable of handling properly. This error increases in magnitude as the cold season approaches, and, when combined with the negative error usually observed over Asia, results in a gradual increase of spurious

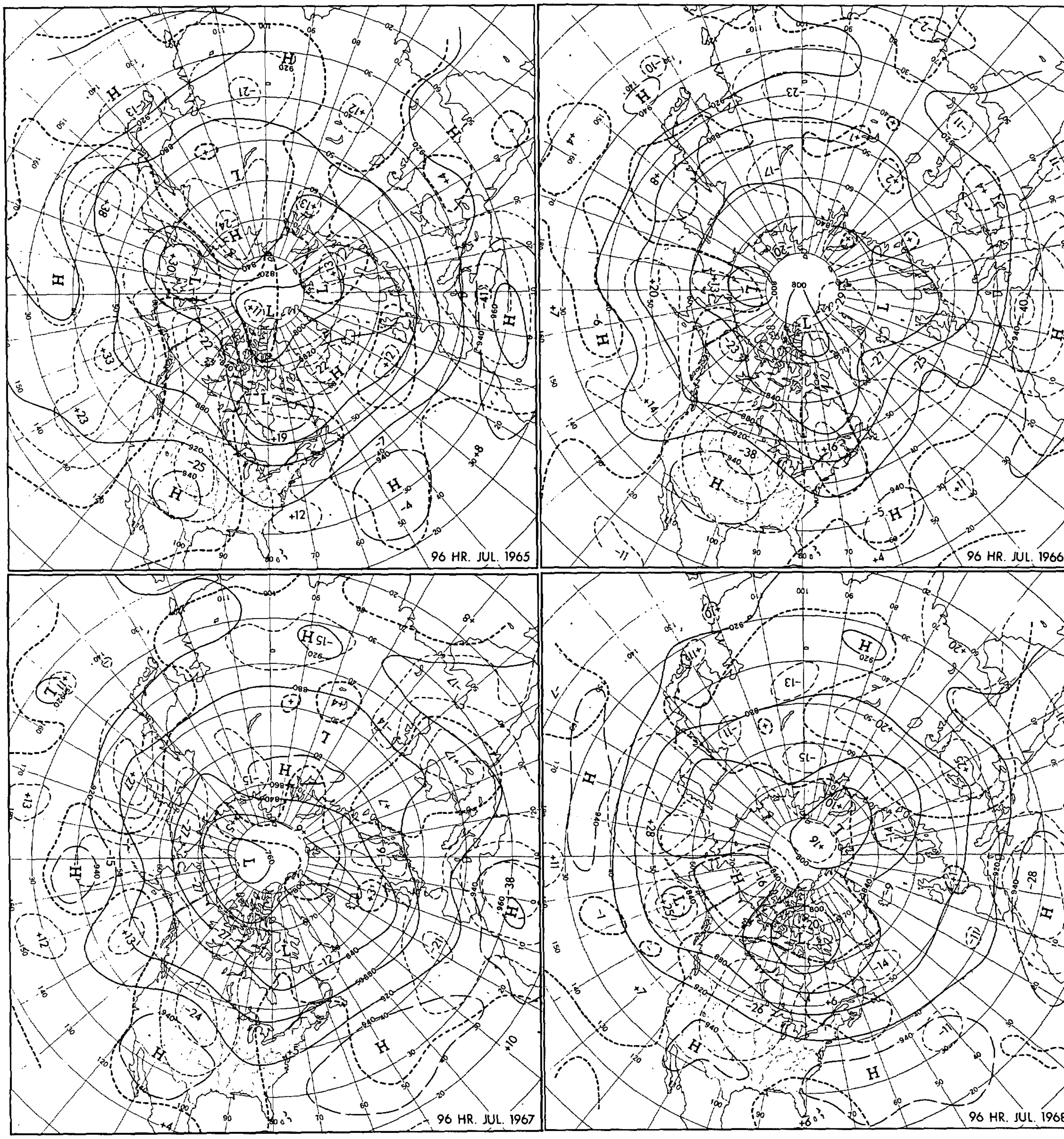

FIGURE 2.—Mean observed 500-mb and 96-hr barotropic error for July 1965, 1966, 1967, and 1968.

southerly flow over the eastern part of the continent. The positive error reaches a peak in January when the midlatitude westerlies have reached their greatest speed and when the ocean-continent thermal contrast and normal heat transfer from the ocean surface to the atmosphere are greatest. The error here is already sizable at 48 hr (fig. 4), but doubles in magnitude by 144 hr (fig. 6). While the size of this error has been gradually reduced from that of the earliest numerical models, it is still very large in extended range barotropic forecasts. Comparison

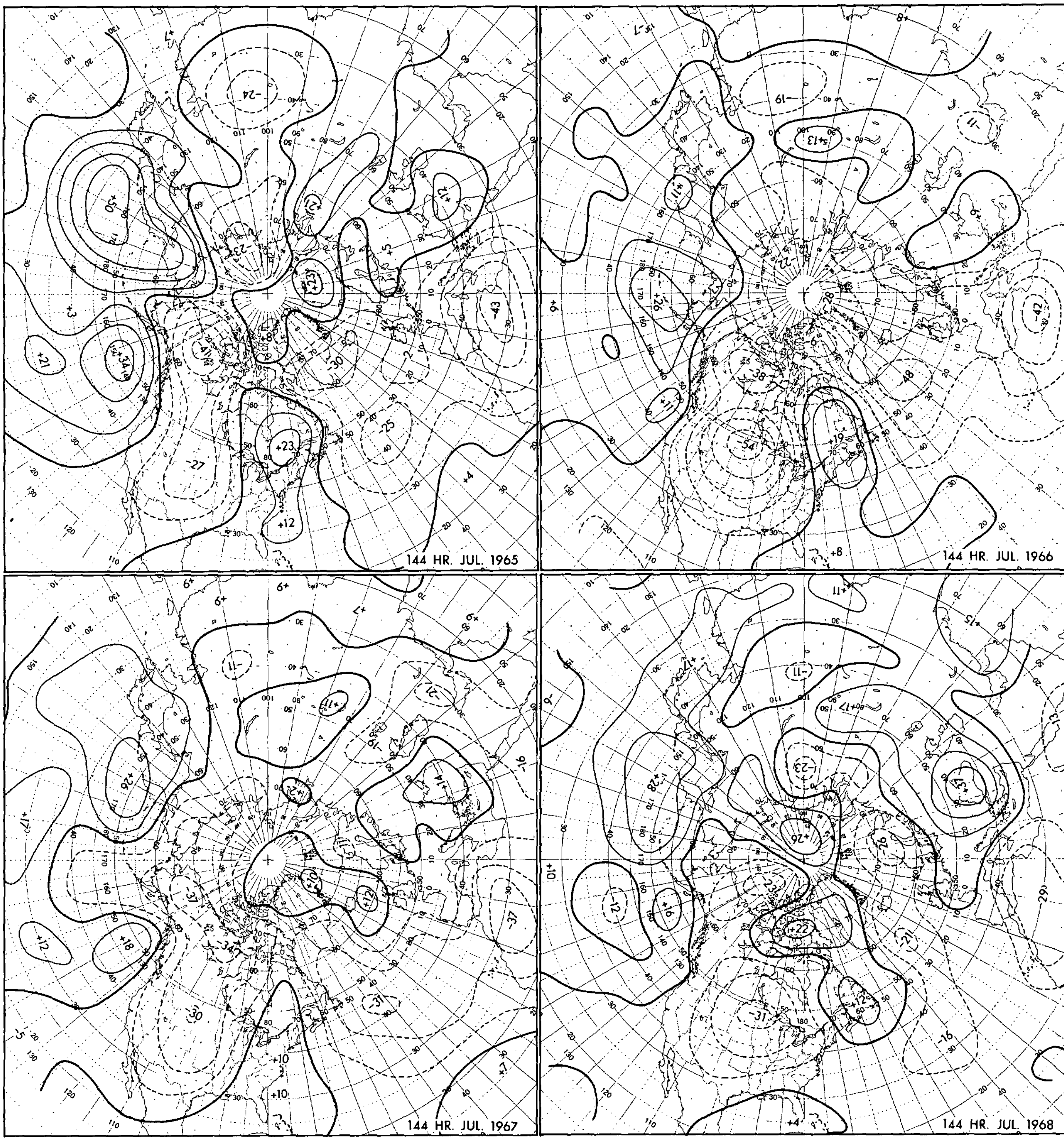

FIGURE 3.—Mean 500-mb 144-hr barotropic error for July 1965, 1966, 1967, and 1968.

of January 1969 error with that of previous Januaries (figs. 4, 5, and 6) shows a marked reduction in magnitude of this error. Some of this decrease may be due to the introduction of latent heat from precipitation into the PE model (February 1967), but part is also related to the position of the polar jet axis which was displaced north of its usual location.

Cyclogenesis is also strongly favored along the east coast of North America by conditions similar to those along the Asiatic coast. Here, however, the average

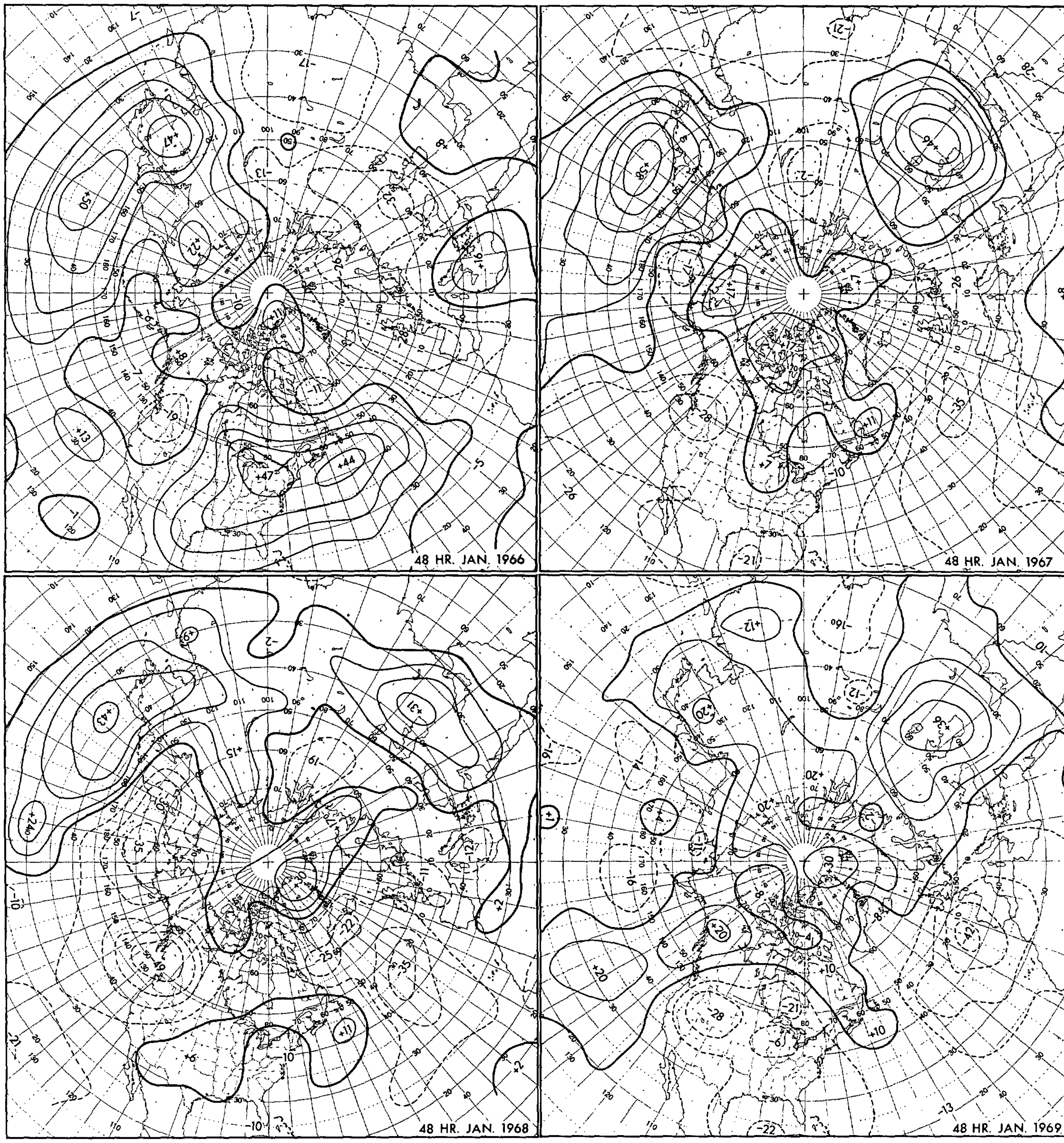

FIGURE 4.—Mean 500-mb 48-hr barotropic error for January 1966 and 1967, and 48-hr PE error for January 1968 and 1969.

positive error in winter is much lower than that in the western Pacific and is usually not well defined until after 48 hr. The error may be less because of smaller thermal contrasts over smaller areas. There appear to be two circulation types which favor the greatest positive errors in the western Atlantic. A large amplitude flow, with a ridge over western North America and a deeper than normal trough near the east coast, such as occurred in February 1968, is associated with large positive errors in the trough. The second type consists of strong blocking over the Atlantic and eastern Canada, as prevailed in January 1966 (fig. 5). At that time, the trough along the coast was not as deep as in February 1968; nevertheless, both circulations had large positive errors of the same magnitude centered near the Middle Atlantic coast and just north of the polar jet axis.

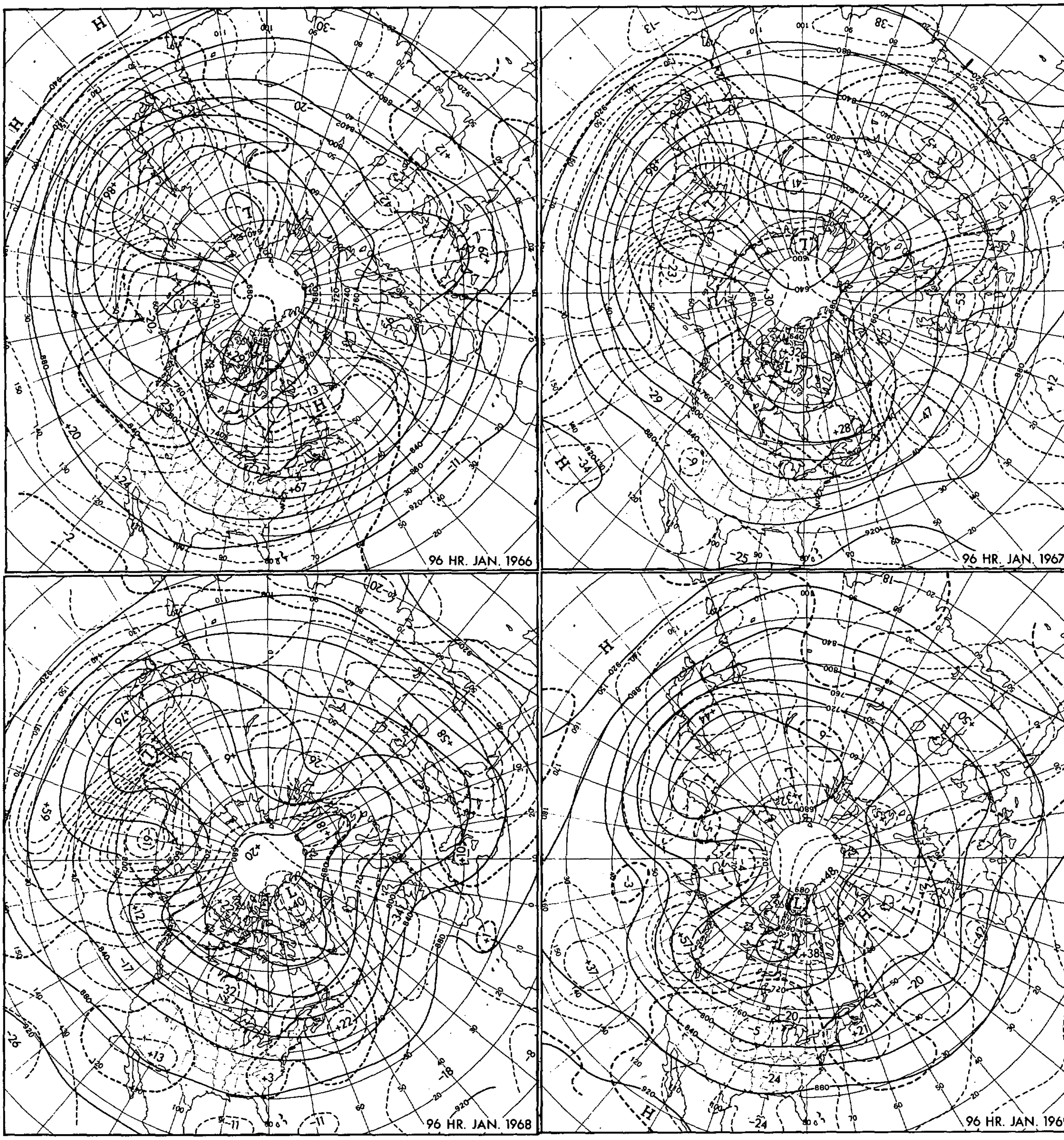

FIGURE 5.—Mean observed 500-mb and 96-hr barotropic error for January 1966, 1967, 1968, and 1969.

There is also a marked preference for positive errors in southeastern Europe and southwestern Asia where planetary troughs are normally found in winter. This error is generally greater than the error in the western Atlantic, but less than that in the western Pacific.

The most persistent negative error in winter is found over the eastern Atlantic and western Europe where ridges dominate the circulation. Negative errors are also observed frequently in the Bering Sea where the largest errors tend to occur with unusually strong mean ridges, as in January 1968 (fig. 5).

SEASONAL ERROR OF *D*+4 FORECASTS

Seasonal mean 700-mb error patterns for the combined PE-barotropic $D+4$ forecasts for 1968 are shown in figure 7. Since only the PE forecasts (to 48 hr) are for the 700-

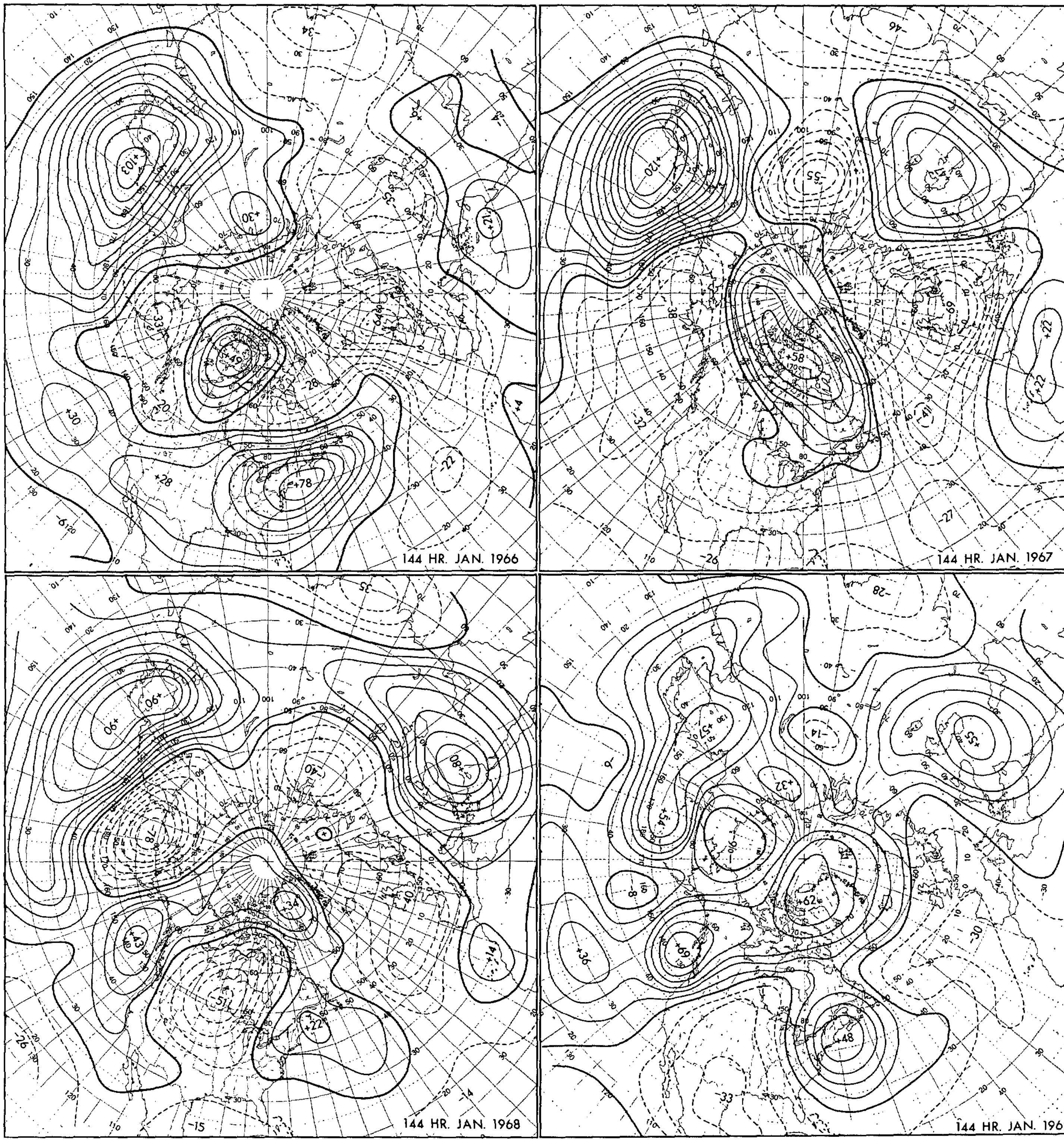

FIGURE 6. Mean 500-mb 144-hr barotropic error for January 1966, 1967, 1968, and 1969.

mb level, it was necessary to reduce the barotropic portion of the $D+4$ forecast from 500 to 700 mb by the use of K factors, the ratio of normal 500- to 700-mb heights.

A comparison of these seasonal errors with their corresponding height anomaly patterns (not shown) shows them to have a good negative correlation. The correlation is very high between the principal centers of positive error and negative height anomaly in mean troughs, with no apparent seasonal differences. A study of shorter period means also shows the forecast error and observed height

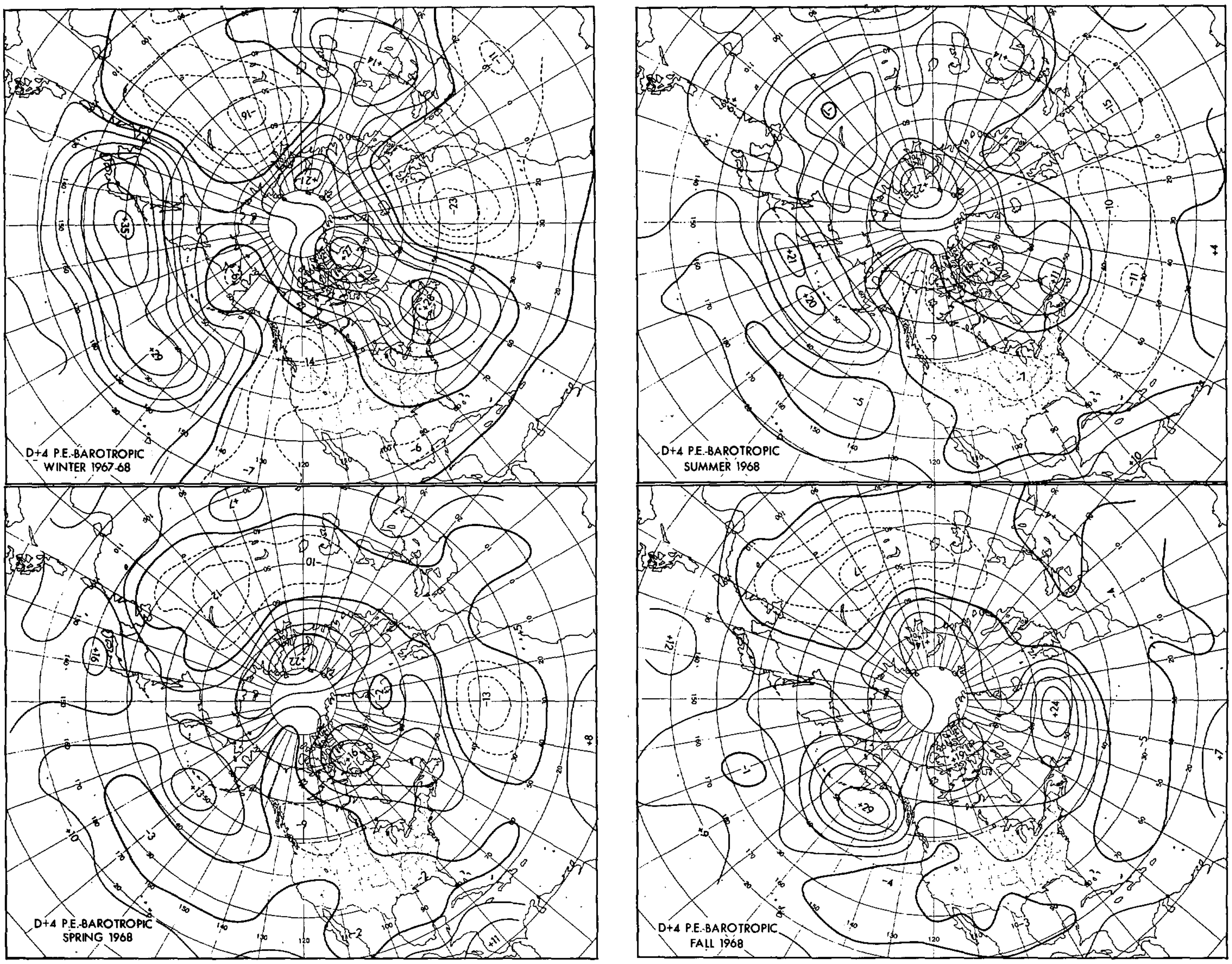

FIGURE 7.—Mean 700-mb $D+4$ PE-barotropic error for winter 1967–1968 and spring, summer, and fall 1968.

anomaly to be highly correlated. Perhaps this good relationship is to be expected since it can be shown by a simple statistical analysis that the variability of the forecast height is less than that of the observed, and that the correlation between these two quantities is less than one.

The seasonal $D+4$ error patterns for 1968 (fig. 7) are similar in many areas to the monthly patterns discussed previously. It is notable that negative error prevailed over western North America in all seasons except fall and over Asia during all seasons except summer. Positive error dominated the polar region and also eastern Canada where troughs are present most of the time.

3. COMPARISON OF SELECTED UPPER LEVEL PE AND BAROTROPIC FORECASTS

FIVE-DAY MEAN $D+4$ ERROR

During 1968 and through May 1969, the PE model was run on Sundays to 156 hr to supplement the regular barotropic output. While some of these runs were not made and others terminated before 156 hr, approximately 20 went the full interval. The $D+4$ error patterns for both the PE and the PE-barotropic forecasts were studied to determine what differences existed between the models.

All cases were examined with respect to high or low zonal index, and it was concluded that differences between the two models were not related to type of circulation. PE heights were almost always lower than barotropic heights at lower latitudes over the Pacific, North America, and the Atlantic, due most likely to lateral boundary problems, with negative error dominant. There was also a stronger tendency for negative error to prevail in the PE forecasts at lower latitudes in summer rather than in winter. This shows up in the PE sea-level (1000-mb) predictions primarily as an underforecast of the strength of the oceanic anticyclones. It was determined that there

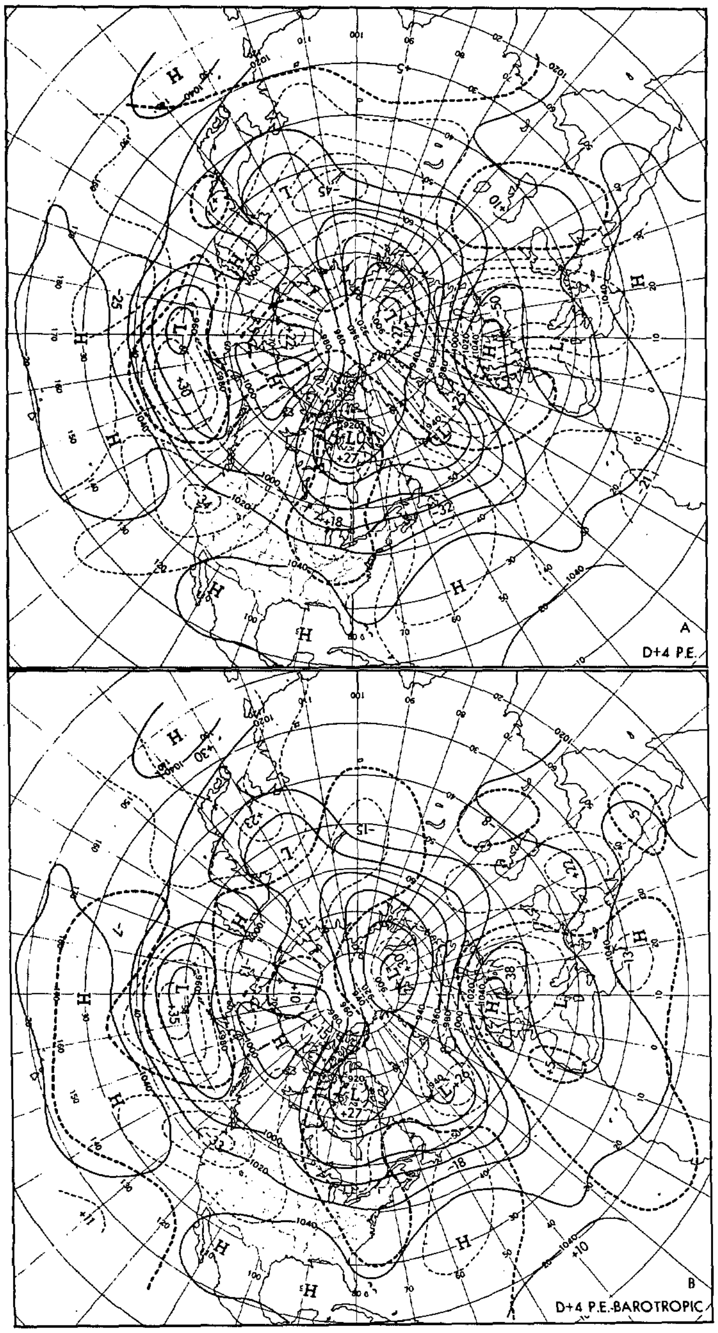

FIGURE 8.—Mean observed 700-mb and $D+4$ error for (A) PE and (B) PE-barotropic models for June 11–15, 1968.

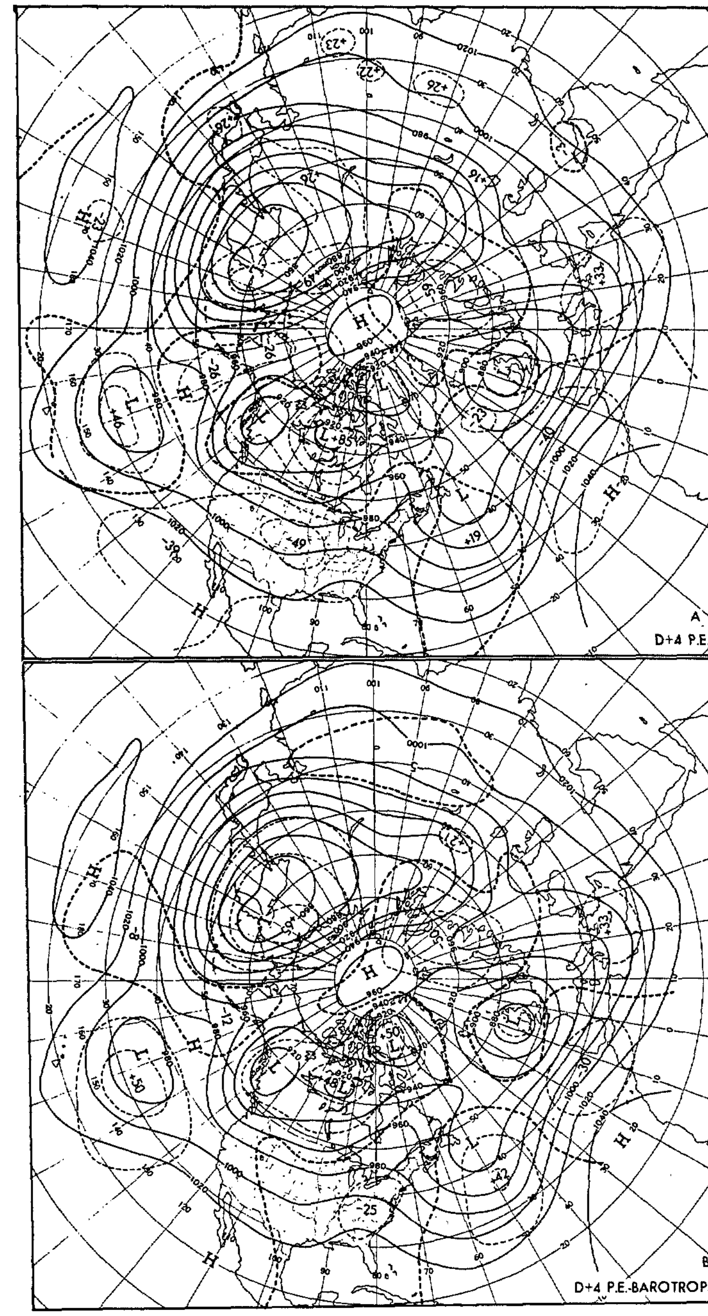

FIGURE 9.—Mean observed 700-mb and $D+4$ error for (A) PE and (B) PE-barotropic models for Jan. 14–18, 1969.

was no tendency for PE heights to be systematically too high at higher latitudes.

Two representative cases of mean 700 mb observed and $D+4$ PE and PE-barotropic error are shown, one for summer (fig. 8) and the other for winter (fig. 9). It is quickly seen that the error patterns in each case are very similar, with differences in the summer case somewhat less than those in the winter case. Note especially that over North America the PE error (fig. 8A) is nearly identical to that of the PE barotropic (fig. 8B). Perhaps this is true generally, since the atmosphere is more barotropic in summer than in winter. In the winter case, the PE error over North America (fig. 9A) was greater than that of the PE barotropic (fig. 9B), with the gradient of error considerably larger. The PE model forecasts the depth of the trough in the western Atlantic more correctly; but in the Pacific, there was little difference in forecasts of the trough intensity near Hawaii.

All $D+4$ forecasts were also compared with the corresponding observed charts. It is concluded that both the PE and PE-barotropic models perform on the average about equally well, and both do a reasonably good job of forecasting the $D+4$ wave pattern. Similar conclusions were reached by Wagner (1967) in his analysis of an earlier test case.

MONTHLY MEAN 96-HOUR ERROR

As part of the Weather Bureau's support program for the serious flood situation that developed in the upper Midwest in the spring of 1969, the NMC prepared during March and April special 3- and 4-day forecasts of temperature anomaly and total precipitation. Guidance for these forecasts was the PE model which was run daily to 108 hr using 0000 GMT data as input. This presented the first opportunity to make a comparison of daily mean monthly error patterns of the PE and barotropic models beyond 48 hr.

In figure 10 is shown the mean observed 500-mb and 96-hr error for (A) PE and (B) barotropic models for 22 cases in March 1969. The PE model was better in predicting trough depths, but there was little difference in the ridge forecasts. In the United States, the gradient of error of the PE model was generally greater than that of the barotropic.

The mean observed 500-mb and 96-hr error for (A) the PE and (B) barotropic models for 13 cases (5-day forecast days) in April 1969 is shown in figure 11. The PE model was generally better than the barotropic over most of North America since the magnitude of negative error (common to both models) was smaller. This was particularly true over western Canada and the eastern United States. It is also noteworthy that the gradient of error over the United States was smaller with the PE model; but in March at 96 hr, the barotropic error gradient was smaller. The PE model in April continued to have smaller positive error than the barotropic in troughs along the Asiatic coast and eastern Mediterranean, but the barotropic forecast of trough intensity in the eastern Pacific and western Atlantic in April was better than that of the PE.

From the limited data presented, it appears that there is very little difference in the large-scale error patterns of the 96-hr 500-mb forecasts from the PE and barotropic models.

4. VERIFICATION OF 5-DAY MEAN UPPER LEVEL CHARTS BY REDUCTION-OF-ERROR SKILL SCORE

In addition to the mean error or bias, another statistic has been used since the winter of 1966–67 to verify the forecasts on a seasonal basis. While this is accomplished by a verification program which produces a large number of statistics, the most comprehensive is the RE skill score. This statistic, devised by Gilman (1969), is derived from "reduction of error" and behaves much like a correlation coefficient, except that it also responds to bias (average error of the sample). It is defined as

$$\text{skill}=\frac{2\overline{FO}}{\overline{O}^2+\overline{F}^2}$$

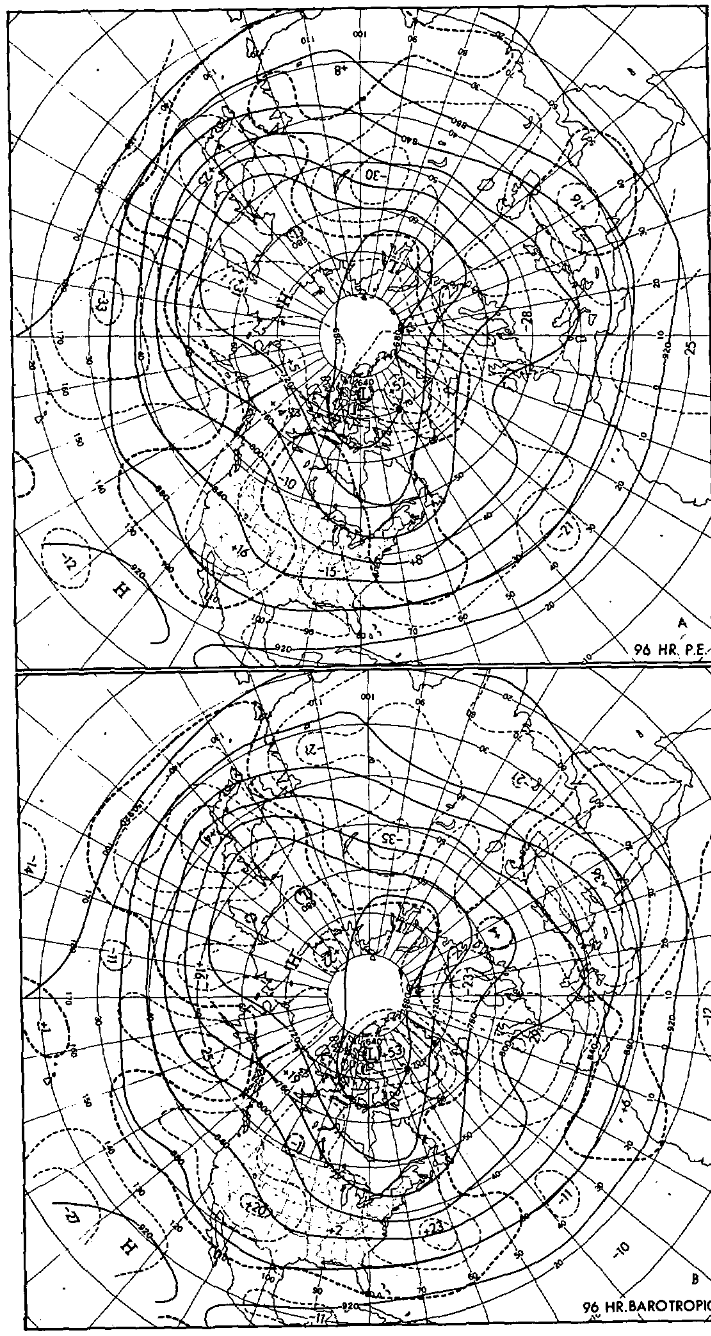

FIGURE 10.—Mean observed 500-mb and 96-hr error for (A) PE and (B) barotropic models for March 1969.

where F and O are the forecast and observed height anomalies, and the bar represents an averaging over the sample of forecasts for each grid point. A perfect score is 100, and a chance score is zero.

Figure 12 shows in graphical form the seasonal RE skill scores for the four operational 5-day mean charts from winter 1966–67 to winter 1968–69. This score is a hemispheric area-weighted average based on a 512-point grid (subset of 1,977-point grid), of all forecasts between latitudes 75° N.–25° N., with persistence of the D_0 and $D+2$ charts used as a predictor of the $D+4$ period.

The $D+2$ has consistently been the best forecast, followed in order by the $D+4$, D_0, and flow charts. The

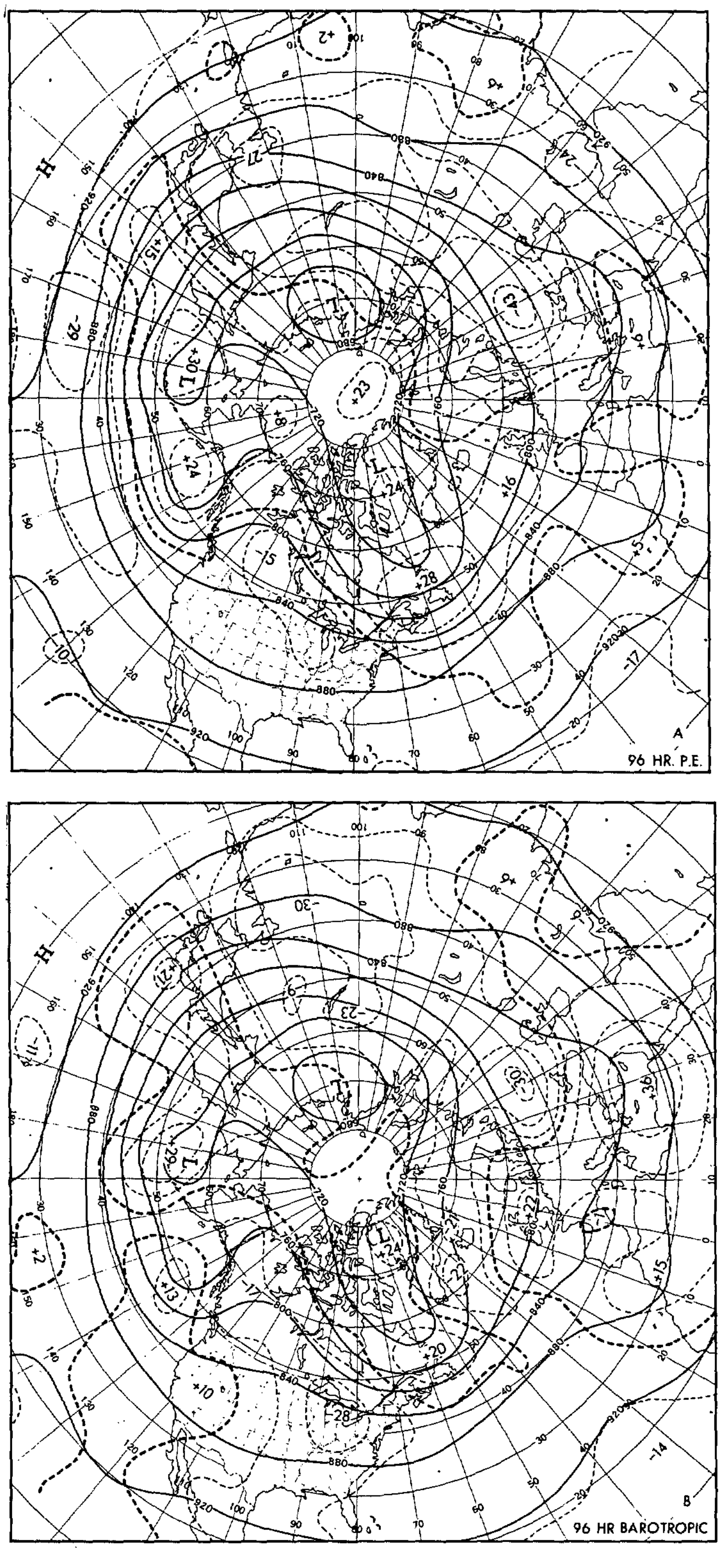

FIGURE 11.—Mean observed 500-mb and 96-hr error for (A) PE and (B) barotropic models for April 1969.

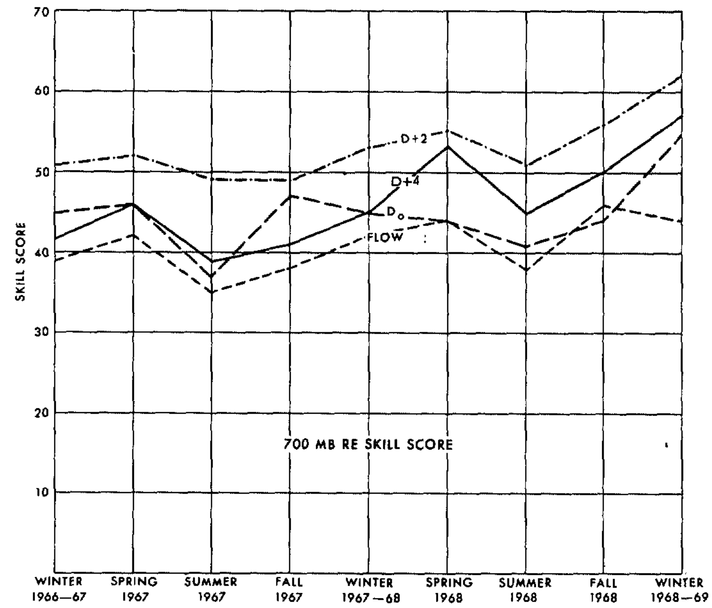

FIGURE 12.—Seasonal mean 700-mb hemispheric RE skill scores (75° N.–25° N.) for D_0, $D+2$, $D+4$, and flow charts for winter 1966–67 through winter 1968–69.

$D+4$ verified about the same as persistence of the D_0 in 1967, but was followed by marked improvement in 1968. The flow chart, on the other hand, displayed less skill than persistence of the D_0 most of the time. All charts showed a gradual increase in skill, with the score for each season higher than that of the previous year. The sharp increase in skill score from fall 1968 to winter 1968–69 for all forecasts except the flow appears to be due primarily to high persistence of the circulation, which was characterized by strong blocking and middle-latitude westerlies south of normal. The only major change made in the PE model during this period was the inclusion of "rough mountains" over western North America. This may have contributed to the increase in skill. RE skill scores for spring 1969 were comparable to those for spring 1968, thus reversing the improvement shown during late 1968 and early 1969.

It should be pointed out here that there is some loss in skill of the PE-barotropic predictions due to reduction from 500 to 700 mb. Verification of forecasts for winter 1966–67 at both levels (Andrews 1967) showed the loss in hemispheric RE skill to be 9 for the $D+4$ and 8 for the flow. The $D+4$ chart then has skill comparable to that of the $D+2$ when verified at 500 mb. Comparison of several observed 5-day mean 500-mb maps reduced to 700 mb with the corresponding observed 700-mb maps showed that the reduction process has a damping affect on the major height anomaly centers. Operational extension of the PE model beyond 48 hr will help eliminate this problem, since 700 mb is one of the six input levels for the PE prediction.

Figure 13 shows meridional profiles of 700-mb hemispheric RE skill score for the same four charts for 1967–1968. All forecasts verified best at high latitudes, worst at low latitudes. The $D+2$ was best at nearly all latitudes, except from 70° N. to 55° N. where the $D+4$ had comparable skill. Comparison of the $D+4$ and flow charts

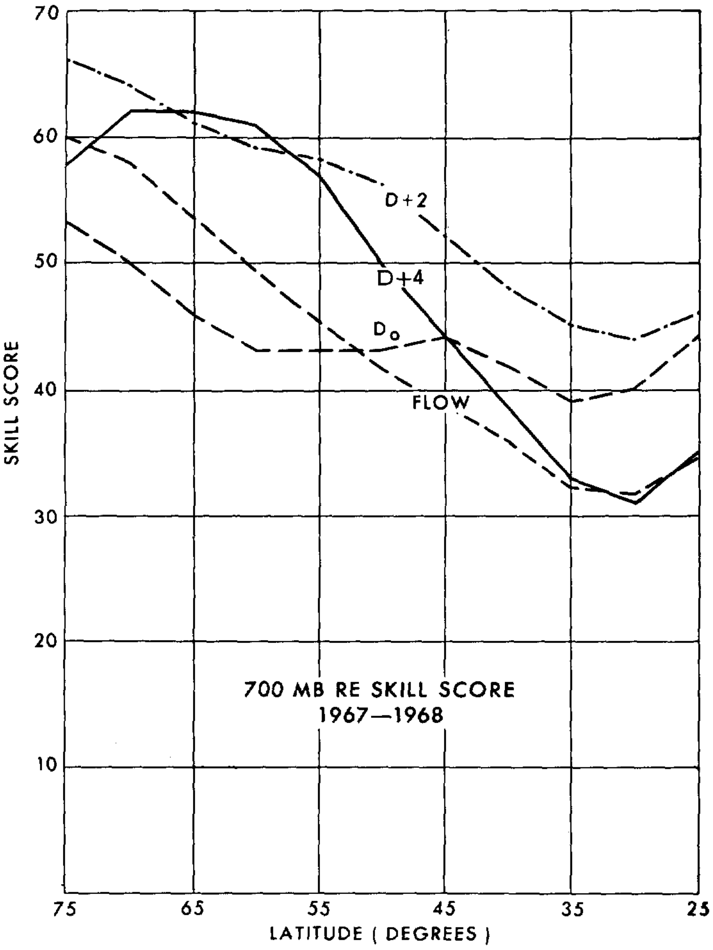

FIGURE 13.—Meridional profiles (75° N.–25° N.) of 700-mb RE skill scores for D_0, $D+2$, $D+4$, and flow charts for December 1966 through November 1968.

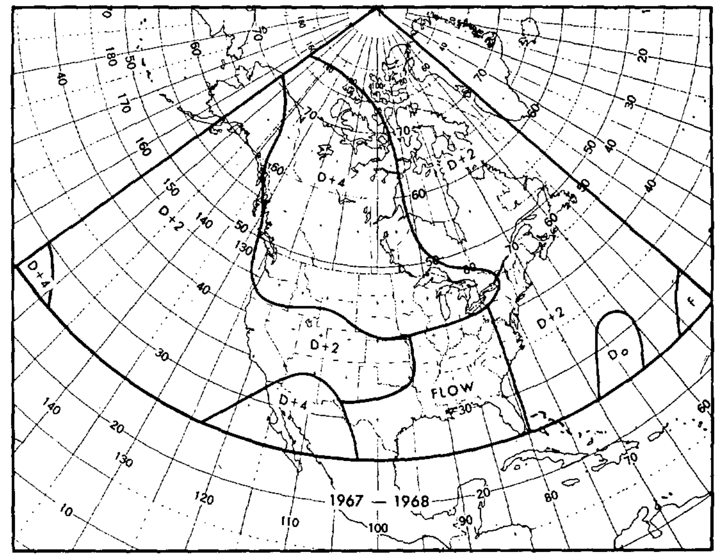

FIGURE 14.—Areas where the indicated chart was better than all of its competitors (as measured by RE skill) for December 1966 through November 1968.

indicates the $D+4$ forecast to be better at all except the extremely high and low latitudes where there was little difference in skill.

Performance of these charts was also compared over North America and adjacent oceans; figure 14 shows those areas where the indicated chart was better than its competitors, as measured by RE skill. It is seen that the $D+2$ was the best forecast over the adjacent oceans, eastern Canada, and the southwestern United States. The $D+4$ forecast was best over western Canada and the northern United States, while the flow chart was the best forecast in the Southeast. An earlier verification for 1961–62 also showed the flow chart to be the best forecast in the Southeast when compared with the D_0 and $D+2$ charts ($D+4$ not available).

5. ELIMINATION OF LARGE-SCALE BIAS

Since the large-scale systematic errors are largely dependent on geographic location and the change in season in some areas, it is possible to remove at least part of this bias from the predicted mean circulation. This was done (post operationally) for the winter of 1966–67 by applying the mean 500-mb $D+4$ error of the previous calendar month to all $D+4$ forecasts of the following month (Andrews 1967). The result for 25 forecasts was an increase in hemispheric RE skill from 51 to 57. Most of this improvement was at middle latitudes and resulted largely from removal of the positive error in the western Pacific. This simple approach could probably be improved by using a running mean error of 3 or 4 weeks ending the day before forecast day, thereby remaining closer to the forecast period. The forecaster would have the option of accepting or rejecting the corrected forecast based on his evaluation of the past and present mean circulation. The pattern of the corrected forecast probably would not differ appreciably from that of the uncorrected one in most areas, but absolute heights could differ significantly.

While the application of a running mean error to the $D+4$ chart might help to improve this forecast, even better results might be obtained if these error patterns were classified by certain key areas. This was attempted previously and summarized in an unpublished report (Andrews 1966) by classifying numerical prediction height errors over North America in summer 1965 by type of initial mean circulation (D_o), using the eastern Pacific as the key area. Composite 5-day mean D_o charts were prepared for those situations which had a deep trough near 150° W. with negative height anomaly near 50° N., 150° W., and for those which had a strong ridge with positive anomaly near the same intersection. These initial circulations were of rather large amplitude, with the ridge cases having a trough along the west coast of North America and negative height anomaly centered over the Northwest. Composite

$D+4$ charts were prepared for these trough and ridge cases and compared with the observed circulations.

The $D+4$ forecast with a deep trough initially in the eastern Pacific proved to be the better of the two predictions over North America, both with respect to pattern and absolute height, with the major ridge-trough system being predicted close to its observed position. The error pattern showed an improvement over the average pattern, as represented by the 96-hr error for July 1965 (fig. 2); and, while similar to that of the ridge cases, had an error gradient of only 120 ft across mid-North America. When the ridge was present, the $D+4$ chart forecast the long-wave features over North America to move eastward much too rapidly, and resulted in an average error gradient of 580 ft.

Examination of similar trough-ridge cases for recent summers (through 1969) indicates that the performance of the combined PE-barotropic model is little different from the situations just described. With more data now available, further stratification may be possible using other key areas and seasons.

The screening technique, so successfully used by Klein (1965) to develop multiple regression equations for the objective prediction of temperature and precipitation using numerically predicted heights as input, might be used to develop similar equations to predict a field of corrected heights. Whether this approach, or the application of a running mean error, would lead to increased skill of the objective temperature and precipitation forecasts is not known, but some improvement seems likely since the equations are so sensitive to slight differences in predicted heights.

6. SUMMARY

Monthly mean daily and 5-day mean upper level errors of the National Meteorological Center's operational numerical forecasts have been shown to be closely related to the mean circulation in such a way that the amplitude of most troughs and ridges is underforecast. The pattern of the daily errors is generally established by 48 hr with a gradual increase in magnitude to 144 hr. The major error centers are highly correlated with the observed height anomalies and also vary in accord with the normal seasonal circulation.

Comparison of daily and 5-day mean forecasts centered 4 days in advance ($D+4$ prognoses), using the operational (combined PE-barotropic) and PE models, indicates that both models perform about equally well over the hemisphere as a whole. The PE model, however, begins to show a definite negative height bias at low latitudes by 96 hr, particularly in summer. This apparently is due to lateral boundary problems and is revealed on the PE sea-level (1000-mb) prognoses, where pressures are forecast too low.

The NMC operational model has shown a gradual but steady improvement over the hemisphere between latitudes 75° N.–25° N. since winter 1966–67. This is indicated by an increase in seasonal reduction of error (RE) skill scores for the 5-day mean forecasts centered both 2 days ($D+2$) and 4 days ($D+4$) in advance.

Removal of the large-scale systematic errors can result in improvement of mean circulation forecasts, as measured by RE skill. Classification of these errors by circulation type may prove to be even better, possibly leading to improved objective predictions of temperature and precipitation.

ACKNOWLEDGMENTS

I wish to express my appreciation to Philip F. Clapp for his many helpful suggestions in the preparation of this report, to Robert B. Danson who did the programming, Roy D. Topper for drafting the figures, and Mrs. Claudia Briner for typing the manuscript.

REFERENCES

Andrews, James F., "Characteristics of Numerical Prediction Errors in Summer," Extended Forecast Division, Weather Bureau, ESSA, Suitland, Md., July **1966**, 11 pp., (unpublished manuscript).

Andrews, James F., "Verification of Selected Upper-Level 5-Day Mean Charts for Winter 1966–67," Extended Forecast Division, Weather Bureau, ESSA, Suitland, Md., Nov. **1967**, 11 pp., (unpublished manuscript).

Budyko, M. I., Editor, *Atlas Teplovogo Balansa Zemnogo Shara*, (Atlas of the Heat Balance of the Earth), Mezhduredomstvennyĭ Geofizicheskiĭ Komitet pri Prezidiume, U.S.S.R. Academy of Sciences, Moscow, **1963**, 69 pp.

Cressman, George P., "A Three-Level Model Suitable for Daily Numerical Forecasting," *Technical Memorandum* No. 22, U.S. Weather Bureau, Washington, D.C., **1963**, 22 pp.

Dunn, Carlos R., "Further Characteristics of the 36-Hr. Prognoses Made With the Operational Barotropic Model," *Technical Note* No. 15, U.S. Weather Bureau, Washington, D.C., Apr. **1964**, 8 pp.

Fawcett, Edwin B., "Systematic Errors in Operational Baroclinic Prognoses at the National Meteorological Center," *Monthly Weather Review*, Vol. 97, No. 9, Sept. **1969**, pp. 670–682.

Gilman, D. L., National Meteorological Center, Weather Bureau, ESSA, Suitland, Md., **1969**, (personal communication).

Klein, William H., "Application of Synoptic Climatology and Short-Range Numerical Prediction to Five-Day Forecasting," *Research Paper* No. 46, U.S. Weather Bureau, Washington, D.C., June **1965**, 109 pp.

Martin, Donald E., "An Investigation of Systematic Errors in the Barotropic Forecasts," *Tellus*, Vol. 10, No. 4, Nov. **1958**, pp. 451–465.

Namias, Jerome, and Collaborators, "Application of Numerical Methods to Extended Forecasting Practices in the U.S. Weather Bureau," *Monthly Weather Review*, Vol. 86, No. 12, Dec. **1958**, pp. 467–476.

Shuman, Frederick G., and Hovermale, John B., "An Operational Six-Layer Primitive Equation Model," *Journal of Applied Meteorology*, Vol. 7, No. 4, Aug., **1968** pp. 525–547.

Stark, L. Paul, "Positions of Monthly Mean Troughs and Ridges in the Northern Hemisphere, 1949–1963," *Monthly Weather Review*, Vol. 93, No. 11, Nov. **1965**, pp. 705–720.

Wagner, A. James, "Objective Numerical Prediction Out to Six Days Using the Primitive Equation Model—A Test Case," *ESSA Technical Memorandum* No. 39, U.S. Department of Commerce, Weather Bureau, Washington, D.C., May **1967**, 19 pp.

[Received September 15, 1969; revised December 8, 1969]